Structural basis for regulation of rhizobial nodulation and symbiosis gene expression by the regulatory protein NoLR

Soon Goo Leea, Hari B. Krishnanb, and Joseph M. Jeza,1

*Department of Biology, Washington University, St. Louis, MO 63130; and bPlant Genetics Research Unit, United States Department of Agriculture-Agricultural Research Service, University of Missouri, Columbia, MO 65211

Edited* by Eva Kondorosi, Hungarian Academy of Sciences, Biological Research Centre, Szeged, Hungary, and approved March 19, 2014 (received for review February 10, 2014)

The symbiosis between rhizobial microbes and host plants involves the coordinated expression of multiple genes, which leads to nodule formation and nitrogen fixation. As part of the transcriptional machinery for nodulation and symbiosis across a range of Rhizobium, NoLR serves as a global regulatory protein. Here, we present the X-ray crystal structures of NoLR in the unliganded form and complexed with two different 22-base pair (bp) double-stranded operator sequences (oligos AT and AA). Structural and biochemical analysis of NoLR reveals protein–DNA interactions with an asymmetric operator site and defines a mechanism for conformational switching of a key residue (Gln56) to accommodate variation in target DNA sequences from diverse rhizobial genes for nodulation and symbiosis. This conformational switching alters the energetic contributions to DNA binding without changes in affinity for the target sequence. Two possible models for the role of NoLR in the regulation of different nodulation and symbiosis genes are proposed. To our knowledge, these studies provide the first structural insight on the regulation of genes involved in the agriculturally and ecologically important symbiosis of microbes and plants that leads to nodule formation and nitrogen fixation.

tenzymatic activator NodD, which is a member of the LysR-type transcriptional reg

The symbiosis between rhizobial bacteria from the Rhizobium, Sinorhizobium, Mesorhizobium, Azorhizobium, and Bradyrhizobium genera and leguminous plants leads to the formation of root nodules (1, 2). These plant organs are specialized for nitrogen fixation and assimilation and are of major ecological and agricultural importance. For example, nitrogen-fixing nodules account for one quarter of total nitrogen fixed globally each year (3). The development of nitrogen-fixing nodules by rhizobia involves a variety of interactions between the plant and microbe; however, at the center of this process are a set of nod (nodulation) genes required for the synthesis of oligosaccharide-nodulation factors, for determining host-plant specificity, and for optimizing the efficiency of symbiosis (4–6). Successful interaction between the rhizobium and host plant requires expression of both positive and negative transcriptional control of genes related to nodulation and symbiosis (7, 8).

Expression of nod genes in rhizobium is regulated by flavonoids released from the host plant in conjunction with the positive activator NodD, which is a member of the LysR-type transcriptional regulator (LTTR) family of proteins (9, 10). Extensive analyses reveal that, in response to small molecules produced by the plant, the rhizobial NodD protein binds to a cis-acting element—the nod box—located upstream of genes required for nodulation (11–16). In addition to the positive control provided by NodD, negative regulation of the nod regulon in Rhizobium melliloti and other rhizobia by NoLR occurs (17, 18).

In rhizobia, NoLR modulates expression of the NodD activator protein, the core nod genes, and multiple genes involved in symbiosis (17, 19–24). Based on amino acid sequence homology, NoLR was proposed to be a helix-turn-helix family member that binds to a nonpalindromic consensus motif (A/T)TTAG-N6-A(T/A) (17).

NoLR is well-conserved across multiple Sinorhizobium and Rhizobium species (23). Originally, NoLR was identified as a negative regulator of nodulation that bound to overlapping transcription initiation sites in the nodD1 and nodA promoters and at the nodD2 promoter (17). Differential regulation of nod genes by NoLR suggested that only the genes related to the synthesis of the core nod factor structure were controlled by this transcription factor (22). Interaction of NoLR with the target DNA site led to reduced levels of nod factors (22). Control of Nod factor production by NoLR may also aid in optimization of nodulation specificity, as NoLR binding sequences were later found in the promoter regions of nodABC, nodD1, tsl-nodD2, noLR, hesB, and nodZ (23). Transcript levels of noLR are high in free-living rhizobia and in the bacteroid but are down-regulated by luteolin, a nod gene inducer (24). Subsequent studies implicated NoLR as a global regulatory factor that responds to environmental factors to fine-tune a range of symbiotic signals, not just the genes required for nodulation, and that the absence or presence of NoLR affects symbiotic interactions with host plants (19, 20, 23, 24). For example, NoLR represses expression of the type II secretion system tisl gene, which is required for secretion of nodulation outer proteins (nops) that are beneficial for symbiosis of Sinorhizobium fredii with some soybean cultivars (24). The molecular basis for NoLR recognition of nonpalindromic DNA target sites and subsequent control of nodulation and symbiosis genes is unclear.

To understand how NoLR functions as a global regulator of nodulation, we used a combination of X-ray crystallography, structural analysis of NoLR, and computational insights to provide an atomic view of how this transcription factor responds to DNA sequence and environmental changes. We investigated the conformational switching of NoLR and its impact on DNA binding and transcriptional efficiency.

Significance

Nitrogen nodules formed by the symbiosis of rhizobial microbes and legume roots are essential for fixation of nitrogen in the environment. As part of the symbiosis that leads to nodule formation, a series of changes in gene expression of the Rhizobium must occur. The protein NoLR is a global regulator of rhizobial genes for symbiosis and nodulation. Here, we describe the three-dimensional structure of this transcription factor in unliganded and DNA bound forms. These structures show how NoLR recognizes asymmetric DNA binding sites and reveal a previously unknown mechanism for conformational switching that alters the energetics of interaction to accommodate variable DNA sequences. Two models for the role of NoLR in the regulation of nodulation and symbiosis genes are also proposed.

Author contributions: S.G.L., H.B.K., and J.M.J. designed research; S.G.L. performed research; S.G.L. and J.M.J. analyzed data; and S.G.L., H.B.K., and J.M.J. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 4OMZ (binding NoLR and oligo AT DNA), 40M2 (unliganded NoLR), and 4ON0 (binding AA DNA)].

1To whom correspondence should be addressed. E-mail: jez@wustl.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1402243111/-/DCSupplemental.
thermodynamic analysis of protein–DNA interactions, and site-directed mutagenesis. The structures of NolR in the unbound form and in complex with two different 22-base pair (bp) double-stranded DNA fragments (oligos AT and AA) reveal a homodimeric protein adopting a winged helix-turn-helix fold. These structures suggest a previously unknown mechanism for conformational switching of a key glutamine side-chain to accommodate DNA-sequence variation in nonpalindromic operator sites without a loss of interaction affinity but with altered thermodynamic contributions to binding. Models for the role of NolR in the regulation of nodulation and symbiosis genes are proposed.

Results

Overall Structure of NolR. NolR from S. fredii USDA191 was expressed as an N-terminally His-tagged protein in Escherichia coli and purified by nickel-affinity and size-exclusion chromatography. The His-tag was removed by thrombin digestion for crystallization. NolR migrated as a dimeric 26-kDa species (monomer M, ~ 13 kDa) by size-exclusion chromatography. Crystals of uncomplexed native NolR and selenomethionine (SeMet)-substituted NolR in complex with a 22-bp oligonucleotide duplex corresponding to the consensus NolR DNA binding motif (oligo AT) were obtained and optimized for data collection (Table S1). The 3D structure of SeMet-substituted NolR bound to oligo AT was determined using single-wavelength anomalous dispersion (SAD) phasing. The resulting model was used to solve the structure of unliganded NolR by molecular replacement.

The overall structure of NolR reveals that the protein is a winged helix-turn-helix transcription factor (25) (Fig. 1 A and B). Two α-helices (α1 and α5) of each monomer form the coiled-coil dimerization interface of NolR. A triangular set of α-helices (α2–α4) positions α3 (residues 45–52) and α4 (residues 55–69) as the helix-turn-helix motif for interaction with the DNA major groove, and the “wing,” a two-stranded antiparallel β-sheet (β1a and β1b), extends into the minor groove (Fig. 1B). For cocrystallization, the conserved NolR promoter site from R. meliloti was used (Fig. 1C) (17). The operator contains two conserved motifs with variable positions that can be either A or T. Comparison of the free and bound forms of NolR indicates that the structure changes little upon DNA binding with a 0.433-Å root-mean-square deviation (rmsd) for 196 Cα atoms in the homodimer. The crystal structure reveals that the NolR dimer binds to position residues of α4 from one monomer to contact the first sequence block (Fig. 1 B and C, purple) on the 5′ strand of the operator and that the same helix of the second monomer interacts with the sequence of the second consensus block (Fig. 1 B and C, red) on the 3′ strand.

Sequence and structural comparisons identify NolR as a member of the ArsRSmtB family of transcription factors (26–29). NolR shares 22–40% amino acid sequence identity with BigR, HlyU, CadC, CtrA, NmtR, and SmtB (Fig. S1A). Moreover, the secondary structure features forming the helix-turn-helix motifs of these proteins are highly conserved (Fig. S1A). A search of the Protein Data Bank using the DALI server (30) identifies BigR, HlyU, and CadC as the closest structural relatives of NolR (Fig. S1 B–D) with Z-scores of 14.4–14.9 and 2.0–2.7 Å rmsd for 95–97 Cα atoms (28, 31–33). The major differences between these proteins occur in the length and positioning of the N-terminal α-helical region at the dimerization interface (Fig. S1 B–D). Residues of the regulatory metal-binding sites found in other ArsRSmtB family members are missing from NolR (26–29, 31–33).

Asymmetric Operator Site Recognition by NolR. The structure of NolR complexed with DNA (Fig. 1 B and C) provides detailed information on how this homodimeric protein recognizes an asymmetric operator site to regulate expression of nodulation and symbiosis genes. Electron density for oligo AT DNA bound to NolR was well-defined (Fig. S2A). NolR binds to the operator, with the α4 helix of each monomer positioned 39 Å apart within major grooves of the DNA duplex (Fig. 1B). The molecular surface of NolR along the DNA-binding interface of each monomer is positively charged and provides an interaction surface for phosphogroups of the DNA whereas the opposite side of NolR is largely negative in charge (Fig. S2B). Analysis of the operator DNA geometry using 3D-DART (34) shows that the duplex bends 16.8° from an ideal B-form upon interaction with NolR (Fig. S2C). Each monomer of the NolR dimer interacts with a sequentially distinct half-site on the operator primarily through residues on α4 (Fig. 2).

Within the first block of the consensus sequence (Fig. 2, purple), two clusters of residues from chain A form extensive interactions with the phosphate backbone of each DNA chain (Fig. 2 A and B). Glm79, Ile81, and Tyr83 hydrogen bond with T1 and A2 of the 5′ strand. On the 3′ strand, Asn28, Arg31, and His62 contact the phosphates of T17 and C18. Glm56, Ser57, Ser60, and Glm61 provide hydrogen-bond interactions with T1, A2, T3, and T4 of the 5′ strand. Ser57 and Glm61 hydrogen bond with T19′ and A20′ of the complementary strand. At the second half site, (Fig. 2C, red) a similar set of residues from monomer B provides additional protein–DNA contacts (Fig. 2 A and C). Gly46, Arg67, Ile81, and Tyr83 of chain B interact with the phosphate groups of C6′, T7′, and T8′ on the 3′ strand. In the second site, Ser60 from α4 hydrogen bonds to the phosphate of T7′ instead of a nucleotide base. Interactions with the 5′-strand phosphates of C12 and C13 are contributed by Asn28, Lys30, Arg31, and His62. The β-sheet wing of monomer B binds in the minor groove to place Glm79, which is on the loop between the two β-strands, within hydrogen-bond distance of A21 and A22 of the 5′ strand (Fig. 2 D and Fig. S3). Ser57 of
monomer B provides bridging contacts with the nucleotide bases of T’ on the 3’ strand and G15 of the 5’ strand. These interactions are similar to those observed in chain A. The side-chain of Gln56 in chain B adopts a conformation that flips the amide group away from T7’ of the 3’ strand. In contrast, the side-chain of Gln56 in the first site is oriented toward the A2 adenine ring.

Movement of Gln56 in Recognition of Variable DNA Sequence. The shift in position of Gln56 in each half-site of the oligo AT DNA, either to interact with A2 in the first site or away from T7’ in the second site, suggested that movement of this residue may play a role in the previously observed recognition by NolR of variable operator sites (17–24). To examine the possible role of Gln56 as a conformational switch in DNA binding, NolR was crystallized with oligo AA DNA (Fig. 3A). The oligonucleotides used to form the duplex maintain the sequence of the first half-site but substitute a T for A17 and an A for T7’ on the 5’ and 3’ strands, respectively. The 3.0-Å resolution crystal structure of NolR in complex with oligo AA was solved by molecular replacement (Table S1). The overall structures of NolR with either DNA bound are nearly identical with a 0.2-Å rmsd for 196 Cα atoms in the homodimer. Comparison of the orientation of Gln56 in the DNA binding sites of each structure show that the interaction with A2 in the first half-site was identical (Fig. 3B and Fig. S4).

In contrast, the side-chain of Gln56 in the second site, which was oriented away from T7’ in the NolR•oligo AT crystal structure, flips to position the amide group for interaction with A7’ in the NolR•oligo AA complex (Fig. 3B and C). Movement of Gln56 to accommodate the variable A/T positions in each half-site of the NolR consensus DNA binding sequence provides a mechanism for recognition of diverse operator sites of genes involved in nodulation and symbiosis of rhizobia and plants (17).

Analysis of DNA Binding by NolR Using Isothermal Titration Calorimetry. The X-ray crystal structures of NolR bound to oligo AT (Fig. 2) and oligo AA (Fig. 3) provide molecular insight on how changes in the position of Gln56 allows for interaction with operator sites with varied sequence at key positions. To examine NolR binding to varied target sites, isothermal titration calorimetry (ITC) was used to characterize interaction with oligos AT and AA. Binding of each DNA duplex to NolR was observed by ITC (Fig. S5); however, the energetic contributions to protein–DNA interaction varied for each sequence (Table 1). Analysis of DNA binding to NolR indicated binding of one dimer per operator site, as observed crystallographically, with comparable affinity ($K_d \sim 0.4 \mu M$). Interestingly, the thermodynamics of interaction were distinct. NolR binding to oligo AT displayed a greater entropic contribution to protein–DNA interaction whereas association with oligo AA was dominated by enthalpic energy. These data indicate that differences in operator site sequence do not alter binding affinity but do change the energetics driving NolR–DNA interaction.

The structures of NolR also implicate key residues as important for DNA binding. To examine the contribution of Arg31, Gln56, Ser57, Ser60, and Gln61, a series of site-directed mutants

---

Lee et al. PNAS Early Edition | 3 of 6
were generated, expressed, and purified. The Q56A mutation was designed to remove the mobile side-chain from NolR for examination of the energetics of protein–DNA interaction by ITC. Arg31, which provides a charge–charge interaction with the DNA backbone, is invariant across members of the ArsR/SmtB family (Fig. S1). The ArsR/SmtB transcription factor family (Fig. S1). Gln61 of NolR is either a glutamine (NolR, BigR, HlyU) or a histidine (CycD, CzaR, MntR, SmtB) in the ArsR/SmtB family (Fig. S1). The ArsR/SmtB family (Fig. S1). Unlike other ArsR/SmtB transcription factors (26–29, 31–33), no global structural

### Discussion

Although the symbiosis between rhizobial microbes and host plants that leads to nitrogen-fixing nodules is an ecologically and agriculturally important process (1–3), the molecular basis underlying the transcriptional regulation of nodulation and symbiosis remains incompletely understood. Nodulation requires induction of *nod* gene expression; however, efficient symbiosis with host plants occurs only when these genes are expressed in an appropriate quantitative, spatial, and temporal pattern and involves both positive and negative regulation (7, 8, 35). Mutations that alter either positive or negative regulation of *nod* genes result in aberrant and delayed nodulation phenotypes.

As part of the transcriptional regulation machinery of nodulation and symbiosis in various Rhizobium, NolR was originally identified as a putative helix-turn-helix family member that bound a nonpalindromic consensus motif in the core *nodABC* gene cluster (17). Later studies broadly implicate NolR as a global regulator of multiple symbiosis-related genes. The crystal structure of NolR is consistent with DNA footprinting studies that mapped the operator site as containing the (A/T)TTAG-N9-A(T/A) consensus operator sites (19, 20, 22–24). Later studies broadly implicate NolR as a global regulator of multiple symbiosis-related genes. The crystal structure of NolR is identified as a putative helix-turn-helix family member that bound

### Table 1. Thermodynamic parameters of DNA binding to wild-type NolR

<table>
<thead>
<tr>
<th>DNA</th>
<th>n</th>
<th>K$_D$ µM</th>
<th>$\Delta G$, kcal mol$^{-1}$</th>
<th>$\Delta H$, kcal mol$^{-1}$</th>
<th>$-T\Delta S$, kcal mol$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>1.02 ± 0.01</td>
<td>0.43 ± 0.06</td>
<td>−8.69 ± 1.11</td>
<td>−3.32 ± 0.04</td>
<td>−5.37</td>
</tr>
<tr>
<td>AA</td>
<td>0.98 ± 0.01</td>
<td>0.36 ± 0.03</td>
<td>−8.78 ± 0.69</td>
<td>−6.67 ± 0.05</td>
<td>−2.11</td>
</tr>
</tbody>
</table>

Titrations were performed using 22-bp DNA duplexes.

The 3D structures of NolR in complex with oligo AT (Fig. 2) and oligo AA (Fig. 3) provide new insight on protein–DNA interactions that regulate nodulation and symbiosis gene expression. Binding contacts with the phosphate backbone are largely contributed by residues on α2, α3, and the β-wing. These contacts likely drive nonspecific association with the operator. Base-specific interactions come from Gln56, Ser57, Ser60, and Gln61 on α4 and Gln79 on the β-wing (Fig. 2 and Fig. S3). The hydrogen-bond network between residues of α4 in each half-site and the DNA duplex shows conserved interactions for operator binding. Ser57 anchors protein interaction by contacting T3 in the first half-site and T8' in the second half-site (Fig. 2). Ser57 also provides hydrogen bonds with A20' in site one and G15 in site two. Gln61 supports the bridging interactions of Ser57 at both sides of the operator. The amide of Gln61 hydrogen bonds with T19' and T14 in the first and second half-sites, respectively. ITC analysis of the NolR S57A and Q61A mutants, which showed a lack of interaction, confirms these residues role in DNA binding. Comparison of NolR operators in *nodD1, nodZ, noIR, nodABC*, and *ttsI* with oligos AT and AA reveals that this central set of interactions would likely be conserved across recognition sites (Table S2). In particular, the high conservation of the first half-site suggests that this site is critical for driving NolR–DNA interaction and that nucleotide variation in the second half-site leads to differences in repression by NolR; however, further studies are required to examine how sequence variations potentially modify NolR binding interactions and/or alter target-gene expression.

In addition to these critical contacts with the core consensus motifs, movement of Gln56 allows for accommodation of variable sequences at A2 and T7' (Fig. 3). The structure of NolR with oligo AT showed that the Gln56 side-chain hydrogen bonds with A2' in the first half-site (Fig. S4) but is flipped away from T7' in the second site (Fig. 3). Substitution of T7' with A7' in oligo AA provides an additional hydrogen bond interaction that orients Gln66 toward the base (Fig. 3). Although NolR binds both operator sequences with comparable affinity, the energetics of protein–DNA interaction differs between the two oligonucleotides (Fig. S5 and Table 1). Interaction with oligo AT displays a larger contribution from entropy compared with oligo AA, but introduction of an additional hydrogen-bond interaction with A7' in oligo AA enhances the enthalpic component of binding. Mutation of Gln66 to an alanine did not significantly alter the affinity of NolR for either oligo AT or oligo AA (Fig. S6 and Table 2) but led to a decreased enthalpic contribution to binding of oligo AA. Movement of Gln66 leads to compensatory energetic effects that maintain the binding affinity of NolR for varied operator sites.

The X-ray crystal structures of NolR in complex with DNA also provide a first view of protein–DNA interaction in the ArsR/SmtB transcription factor family (Fig. S1). Unlike other ArsR/SmtB transcription factors (26–29, 31–33), no global structural

### Table 2. Thermodynamic parameters of DNA binding to the NolR Q56A mutant

<table>
<thead>
<tr>
<th>DNA</th>
<th>n</th>
<th>K$_D$ µM</th>
<th>$\Delta G$, kcal mol$^{-1}$</th>
<th>$\Delta H$, kcal mol$^{-1}$</th>
<th>$-T\Delta S$, kcal mol$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>1.00 ± 0.01</td>
<td>0.26 ± 0.01</td>
<td>−8.99 ± 0.46</td>
<td>−3.33 ± 0.01</td>
<td>−5.66</td>
</tr>
<tr>
<td>AA</td>
<td>1.03 ± 0.01</td>
<td>0.82 ± 0.11</td>
<td>−8.30 ± 1.09</td>
<td>−3.77 ± 0.07</td>
<td>−4.53</td>
</tr>
</tbody>
</table>

Titrations were performed using 22-bp DNA duplexes.
changes in the unliganded and DNA bound forms of NolR were observed, which likely results from a lack of metal-binding sites in the dimer interface of NolR. In the ArsR/SmtB proteins, metal binding triggers conformational changes that result in derepression of gene expression by driving repositioning of the helix-turn-helix DNA interaction motif (26–29, 31–33). This movement results in a switch from a “closed” DNA-binding structure to an “open” low-affinity conformation of the homodimer. To date, there is no evidence for other ligand interactions with NolR that trigger conformational changes that alter binding of the operator site.

Previous structural studies of ArsR/SmtB proteins largely focused on the role of metal-dependent conformational changes and not protein–DNA interaction, with one exception. NMR and mutagenesis probed residues in CzrA from *Staphylococcus aureus* for roles in binding to the palindromic 28-bp operator (28). This work suggests that CzrA interaction with target DNA increases protein motion in the allosteric sites and showed essential roles for residues analogous to Gln56, Ser57, Ser60, and Gln61 in DNA binding by CzrA, but did not reveal the details of protein–DNA contacts required for phosphate backbone interactions and/or base specificity.

At the molecular level, interaction of NolR with operator sites of genes related to nodulation and symbiosis suggests at least two models for how the repressor functions. The first model of NolR regulation of gene expression is that of binding to an operator within the transcription initiation site of a target gene (Fig. 4B). The NolR binding sites for nodD1, nodZZ, and nodR (Table S2) are all examples of this arrangement. NodD1 is the major transcription factor for nodulation and nodZ is an α,β-fucosyltransferase for nod-factor assembly (7–9, 14, 35–40). In each gene, a NolR operator is positioned 15–60 bp upstream of the coding region (23). Likewise, the presence of a NolR binding site 35 bp in front of the *nolR* gene suggests that levels of NolR modulate its own expression. Increased expression of NolR in free-living Rhizobia down-regulates expression of nodD1 and other nodulation/symbiosis genes, such as nodZ (22). In this arrangement, NolR binding to the operator would result in repression of gene expression by driving repositioning of the DNA backbone (28). This structure either rigidifies the DNA to alter interaction with the promoter or changes the structure of the nod box to modulate NodD interaction. Further studies are needed to fully understand how these opposing transcription factors modulate gene expression.

In conclusion, the structural and thermodynamic studies presented here provide, to our knowledge, the first insights on the molecular foundation for the regulation of gene expression in nodulation and symbiosis and suggest how the global regulatory protein NolR recognizes variable asymmetric operator sites in the promoter regions of diverse rhizobial genes.

**Methods**

**Protein Expression, Purification, and Mutagenesis.** The coding region of *nolR* was PCR-amplified from *S. fredii* USDA191 genomic DNA using oligonucleotide primers that included *NdeI* and *XhoI* restriction sites, respectively, to facilitate cloning. The PCR product was digested with *NdeI* and *XhoI* and ligated into pET-28a. The resulting vector was then transformed into *E. coli* BL21 (DE3) cells. Transformed *E. coli* cells were grown at 37 °C in Terrific broth containing 50 μg·mL⁻¹ kanamycin until *A*₆₀₀nm = 0.9–0.9. After induction with 0.1 mM isopropyl-β-D-galactopyranoside, the cultures were harvested at 20 °C overnight. Following centrifugation, the cell pellet was suspended in 50 mM Tris (pH 8.0), 500 mM NaCl, 1 mM β-mercaptoethanol (βME), 10% (vol/vol) glycerol and 1% Tween 20. Sonication was used for cell lysis. After centrifugation, the supernatant was passed over an Ni²⁺-nitrilotriacetic acid column. The column was washed with buffer minus Tween 20, and the bound His-tagged protein was eluted using 250 mM imidazole in wash buffer. The N-terminal His-tag was removed by thrombin digestion in dialysis against wash buffer. A mixed Ni²⁺-nitrilotriacetic acidbenzamidine Sepharose column was used to remove undigested protein and thrombin. Size-exclusion chromatography was performed using a Superdex-75 26/60 FPLC column equilibrated in 150 mM Tris (pH 8.5), 100 mM NaCl, 1 mM βME. Fractions corresponding to the protein peak were pooled and concentrated to 10 mg·mL⁻¹ by centrifugal filtration. Protein concentration was determined using the Bradford method, with BSA as a standard. SemMet-substituted NolR protein was produced by inhibition of the *E. coli* methionine biosynthesis pathway (44) and was purified as described above. Site-directed mutants of NolR (R31A, Q56A, S57A, S60A, and Q61A) were generated using the QuikChange PCR method and were expressed and purified using the same methods used with wild-type NolR.

**Protein Crystallography.** All NolR protein crystals were grown by the hanging drop vapour diffusion method at 4 °C. For crystallization of NolR in the presence of oligo AT, two 22-bp-long oligonucleotides (5′-d(TTTAGAGAAGGAGGAGGAG)-3′ and 5′-d(GAGGCTCCCGGCTCCGAAT-3′) were suspended in purification buffer and annealed. Crystals of SemMet-substituted NolR (10.0 mg·mL⁻¹) in complex with oligo AT DNA (1 mM) grew in drops containing 0.1 M NaCl, 0.1 M sodium phosphate monobasic/0.4 M potassium phosphate dibasic, 0.1 M sodium phosphate dibasic/citric acid (pH 4.0). Crystals of native NolR (10 mg·mL⁻¹) were formed in drops of a 1:1 mixture of protein and crystallization buffer [1.6 M sodium phosphate monobasic/0.4 M potassium phosphate dibasic, 0.2 M sodium phosphate dibasic/citric acid (pH 4.2)]. Crystals of native NolR (10 mg·mL⁻¹) in complex with 1 mM oligo AA DNA (1 mM) (5′-d(TTTAGAGAAGGAGGAGGAG)-3′ and 5′-d(GAGGCTCCCGGCTCCGAAT-3′) were obtained under conditions similar to that of the NolR oligo AT complex. All crystals were stabilized in crystallization solution with 30% glycerol before flash freezing in liquid nitrogen for data collection at 100 K. All X-ray diffraction data (wavelength = 0.979 Å) were collected at
beamine 19-ID of the Argonne National Laboratory Advanced Photon Source. HKL3000 (43) was used to index, integrate, and scale diffraction data. The structure of Smet-substituted NolR in complex with oligo AT DNA was determined by SAD phasing. SHELXL (46) was used to determine SeMet positions and to estimate initial phases from the peak wavelength dataset. Refinement of SeMet positions and parameters was performed with MLPHARE (47). Solvent flattening using density modification implemented with ARPH/wARP (48) was used to build an initial model. Subsequent iterative rounds of manual model building and refinement, which included translation-libration-rotation (TLR) refinement, were carried out using COOT (49) and PHENIX (50), respectively. The structures of unliganded NolR and NolR complexed with oligo AA DNA were solved by molecular replacement in PHASER (51) using the SeMet-substituted NolR structure as a search model, with model building and refinement performed with COOT and PHENIX. Water molecules were added to the unliganded NolR model using default parameters in PHENIX. Crystallographic statistics are summarized in Table S1.

**Isothermal Titration Calorimetry**

NolR protein was dialyzed overnight in 150 mM Tris (pH 8.5), 100 mM NaCl, 5 mM MgCl₂, 5% glycerol at 4 °C. Synthetic double-stranded DNA (oligo AT, AA, and TT) were prepared in the same buffer. ITC experiments were performed using a VP-ITC calorimeter (Microcal, Northampton, MA) at 4 °C. Data obtained from the titrations were analyzed using a single-binding site model: ΔH = nKₜₐₜDNAΔH₀/(1 + xKₛ), in which ΔH is the total heat after the i-th injection, Vᵢ is the calorigraphic cell volume, Mᵢ is the concentration of protein in the cell after the i-th injection, ΔH is the corresponding enthalpy change for NolR-DNA binding, and x is the number of nucleotide binding sites, and Kₛ is the equilibrium constant. Estimates of Kₛ and ΔH were obtained by fitting the experimental data using Origin software (Microcal). Values for the change in free energy (∆G) were calculated using ∆G = -RTln(Kₛ), where R is the gas constant (1.9872 cal·K⁻¹·mol⁻¹) and T is absolute temperature.

**Changes in entropy (∆S)** were calculated using ∆S = ∆H - T·∆S, which yields Kₛ = 1/(T·∆S).

**Acknowledgments**

 Portions of this research were carried out at the Argonne National Laboratory Structural Biology Center of the Advanced Photon Source, a national user facility operated by the University of Chicago for the Department of Energy Office of Biological and Environmental Research (DE-AC02-06CH11357).